(1) presenting a collection of training data comprising examples of historical policy and claims data; and

(2) generating on the basis of the training data a plurality of segment models, that together comprise an overall model, wherein each segment model comprises a statistical model of insurance risk that is associated with a specific segment of the training data, said generating comprising:

B1 Cont'd

a) generating alternative pluralities of segment models in one of a top-down fashion and a bottom-up fashion;

b) comparing said alternative pluralities of segment models based on the corresponding statistical models of insurance risk; and

c) selecting a final plurality of segment models and associated segments from among the alternates generated so as to optimize aggregate numerical criteria for the plurality.

REMARKS

Entry of this Amendment is proper under 37 CFR §1.116 since no new claims are added, since the only claim amendment is not in any way related to the reference cited by the Examiner and, therefore, raises no new issues relative to the present rejection, and since entry clarifies for appeal the misinterpretation of terminology in the rejection of record.

Attached hereto is a marked up version of the changes made in the specification and claims by the current Amendment. The attached page is captioned "**Version with markings to show changes made.**"

It is noted that the claim amendments herein are intended solely to more particularly point out the present invention for the Examiner, and not for distinguishing over the prior art or the statutory requirements directed to patentability.

It is further noted that, notwithstanding any claim amendments made herein, Applicant's intent is to encompass equivalents of all claim elements, even if amended herein or later during prosecution.

Claims 1-5 are all of the claims pending in the present Application, and all five claims stand rejected under 35 USC §103(a) as unpatentable over US Patent 5,970,464 to Apte et al., further in view of US Patent Application Publication No. US 2001/0020229 to Lash.

These rejections are respectfully traversed in view of the following discussion.

I. THE CLAIMED INVENTION

As described and claimed (e.g., by independent claim 1), the present invention is directed to a program storage device readable by a machine for constructing segmentation-based models that satisfy constraints on the statistical properties of the segments, including presenting a collection of training data records comprising examples of input values that are available to the model together with the corresponding desired output value(s) that the model is intended to predict generating on the basis of the training data a plurality of segment models, that together comprise an overall model, wherein each segment model is associated with a specific segment of the training data.

The generating includes performing an optimization that includes generating alternate training data segments and associated segment models, evaluating at least one generated segment to determine whether it satisfies at least one statistical constraint, and selecting a final plurality of segment models and associated segments from among the alternates evaluated that have satisfactory evaluations.

An advantage of the present invention is that it provides a technique analogous to the distinction between closed-loop and open-loop systems. That is, as discussed beginning on page 33, beginning at line 14, relative to techniques discussed in the specification, the present invention applies statistical constraints as segment splits are being constructed, thereby guiding the segment construction process.

II. THE PRIOR ART REJECTION

The Examiner asserts that the Apte et al. reference teaches the present invention, but concedes that the primary reference "fails to expressly disclose 'a program storage device readable by a machine tangibly embodying a program of instructions' in the preamble."

To overcome this deficiency, the Examiner then introduces US Patent Application Publication No. US 2001/0020229 to Lash. The Examiner further asserts that one of ordinary skill in the art would have been motivated "to expand Apte's computer-implemented method of underwriting profitability analysis to include a program storage unit, which may incorporate one or more conventional storage devices adapted to read programming data, as taught by Lash, with the motivation of providing means for storage and retrieval of program data and instruction to be used at a later time".

Applicant respectfully disagrees that one of ordinary skill in the art would agree with the Examiner that Apte can reasonably be considered as teaching the concepts of the present invention.

First, it is noted that the rejection of record is understood as containing a typographical error by referring to the Apte et al. reference as being US Patent 5,890,129, rather than 5,970,464.

Second, the present invention provides a technique analogous to a closed loop, as explained briefly in lines 20-25 of page 33, by applying statistical constraints as an integral part of the method *for splitting larger segments into smaller segments*. Thus, even if Apte did automatically execute a fine tuning process, which the Examiner considers as being obvious to do, it would still fail to achieve the technique of the present invention, as described exemplarily by the limitation in claim 1: "evaluating at least one generated segment to determine whether it satisfies at least one statistical constraint".

That is, Applicant respectfully submits that the Examiner incorrectly interprets the lexicon used in the present invention as compared to the lexicon used in Apte, as one of ordinary skill in the art would understand this terminology.

Relative to claim 1, the Examiner makes several incorrect interpretations. On page 3 of the Office Action dated April 23, 2002, the Examiner states that "pure premium

characteristics" is interpreted as meaning "desired output values that the model is intended to predict".

However, the phrase "pure premium characteristics", as used in Apte at lines 20-33 of column 3, actually refers back to the definition in lines 30-33 of column 1, "the premium at which their expected claims payout equals premiums charged". As is well known to actuaries, the pure premium of a risk group for property and casualty insurance can be equivalently estimated either as the product of estimated claim frequency times estimated claim severity (see present application, page 8, lines 28-29), or as the ratio of the sum of the historical claims filed by a risk group divided by the sum of the historical earned exposures of the insurance policies of the risk group, where the earned exposure of a policy is the length of time that the policy has been in force (see present application, page 51, line 7).

The phrase "desired output value(s) that the model is intended to predict", as used in the present invention, corresponds in Apte (column 3, lines 57-58) to "a pair of fields (known as the 'response' variables) will be the target of the data mining process".

Note, however, that the former is intentionally broader in scope than the latter in that the former can refer to one or more response variables, and these response variables need not necessarily correspond to either claim frequency or claim response, as in Apte, column 3, line 59.

Next, on page 4 of the Office Action dated April 23, 2002, the Examiner interprets "actual pure premium" as a "statistical constraint".

The phrase "actual pure premium", as used in Apte, column 3, lines 20-33, refers to "estimated quarterly pure premium", as demonstrated in the illustrative example that immediately follows the use of the phrase "actual pure premium" (Apte, column 4, lines 1-5).

In reality, the phrase "statistical constraint", as used in the present application, has no counterpart in Apte, nor in any prior method for constructing segmentation-based models that Applicant is aware of (with the possible exception of constraints on the number of training data records of various types that must fall into each segments, which exception is not a "statistical constraint" as would be commonly understood in the art). This use of applying statistical constraints to guide the process of splitting larger segments into smaller suitable

population segments is a distinguishing feature of the present invention and provides a "closed loop" technique for generating segments that is heretofore unknown in the art.

Indeed, the technique described at lines 34-37 of column 4 of Apte would be understood as deleting segments that were previously identified by the process described at lines 23-27 of column 7, rather than generating new segments, let alone generating segments as guided by a statistical evaluation.

Next, in the penultimate paragraph of page 5 of the Office Action dated April 23, 2002, the Examiner interprets "fine tuning the eligibility criteria for the product, until the segments that are dragging the overall costs down are satisfactorily removed" as "selecting a final plurality of segments that have satisfactory evaluations".

However, this phrase, as used in Apte, column 4, lines 33-36, refers to the scenario analysis method disclosed in Apte that is specified in detail beginning at column 8, line 56.

It is noted that this scenario analysis in Apte utilizes a segmentation-based model that is produced as a result of data mining. Hence, the scenario analysis must necessarily be performed after data mining is performed. That is, the data mining process disclosed in Apte analyzes historical policy and claims data in order to segment policyholders into homogeneous risk groups. Scenario analysis then provides a method for adjusting the eligibility criteria for an insurance product so as to exclude those risk groups identified in the segmentation-based model whose pure premiums are too high relative to the actual (or anticipated) premium (to be) charged for the insurance product.

In contrast, the phrase "selecting a final plurality of segments that have satisfactory evaluations", as used in the present application, refers to a method *step in a data mining process* that could be used, for example, to segment policyholders into homogeneous risk groups *prior to* performing scenario analysis. As recited in claim 1, in the present invention, segments are evaluated by identifying those segments that satisfy desired statistical constraints.

Relative to claim 2, the Examiner repeats some of the above incorrect interpretations, and additionally makes the following incorrect interpretation:

Apte, column 4, lines 28-39, the Examiner interprets "actual pure premiums" as a "statistical constraint." The Examiner also interprets the generating as being done in a

"closed loop fashion," because the statistical constraint, "actual pure premium", is part of the eligibility criteria which is evaluate and then used to regulate the construction of potential segments (Apte, column 4, lines 8-16).

With regard to the first part, the misinterpretation of "actual pure premiums" as a "statistical constraint" has already been discussed above. With regard to the second part, the Examiner is incorrectly equating scenario analysis in Apte with the construction of segmentation-based models in the present application.

As discussed above, in Apte, scenario analysis is a post-processing step that takes a segmentation-based model as input in order to adjust the eligibility criteria for insurance products. The present application concerns the data mining step that must be performed prior to scenario analysis in order to produce a suitable segmentation-based model. Although the Examiner is correct in implying that scenario analysis is a closed -loop process, scenario analysis has a different purpose than, and cannot be equated with, the data mining process that is the subject matter of the present application.

Relative to claim 3, the Examiner repeats some of the above incorrect interpretations, and additionally makes the following incorrect interpretation:

Apte, column 4, lines 28-39; the Examiner interprets ... "Fine tuning" as a form of "adjusting".

The Examiner is again incorrectly equating scenario analysis in Apte with the construction of segmentation-based models in the present Application.

Hence, turning to the clear language of the claims, there is no teaching or suggestion in the Apte reference of "... generating alternate training data segments and associated segment models; evaluating at least one generated segment to determine whether it satisfies at least one statistical constraint ...", as required by claim 1. Similar language is in claims 2-4. Lash does nothing to make up for this deficiency.

Accordingly, for this reason alone, claims 1-5 are fully patentable over the Apte reference.

Further, the other prior art of record has been reviewed, but it too even in combination with the Apte or Lash references, either alone or in combination, fails to teach or suggest the claimed invention.

III. FORMAL MATTERS AND CONCLUSION

In view of the foregoing, Applicant submits that claims 1-5, all the claims presently pending in the application, are patentably distinct over the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary in a telephonic or personal interview.

The Commissioner is hereby authorized to charge any deficiency in fees or to credit any overpayment in fees to Assignee's Deposit Account No. 50-0510.

Respectfully Submitted,

Date: 12/30/02

Frederick E. Cooperrider
Reg. No. 36,769

McGinn & Gibb, PLLC
8321 Old Courthouse Road, Suite 200
Vienna, Virginia 22182
(703) 761-4100
Customer No. 21254

CERTIFICATION OF TRANSMISSION

I certify that I transmitted this Amendment under 37 CFR §1.116 to Examiner M. Kapadia on December 30, 2002, via facsimile to (703) 305-7687 (organization fax number listed in Office Action) and to (703) 872-9327 (official number listed in USPTO web page for TC 3600).

Frederick E. Cooperrider # 36,769
Reg. No. 36,769 12/30/02

S/A 09/302,154
Docket: YO999-214

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE CLAIMS:
Claim 4 has been amended, as follows:

4. (Amended) A program storage device readable by a machine, tangibly embodying a program of instructions executable by the machine to perform a method for constructing segmentation-based models of insurance risks, the method comprising:

(1) presenting a collection of training data comprising examples of historical policy and claims data; and

(2) generating on the basis of the training data a plurality of segment models, that together comprise an overall model, wherein each segment model comprises a statistical model of insurance risk that is associated with a specific segment of the training data, said generating comprising [performing optimization comprising]:

a) generating [alternate training data segments and associated segment models] alternative pluralities of segment models in one of a top-down fashion and a bottom-up fashion;

b) [evaluating the generated segment models using numerical criteria derived from statistical models used by actuaries to model insurance risks,] comparing said alternative pluralities of segment models based on the corresponding statistical models of insurance risk; and

c) selecting a final plurality of segment models and associated segments from among the alternates generated so as to optimize aggregate numerical criteria for the plurality.